IReNe: Supplemental Materials

1. Quantitative results

We include a a detailed description of the quantitative results for each of the scenes of all the three datasets that we used in our work Synthetic NeRF [5], LLFF [4] and Mip NeRF 360 dataset [1]. In this more detailed description of our numerical experiments we seek to provide the reviewers and readers an in-depth insight at how our method performs with respect to the other state of the art approaches with whom we compared, PaletteNeRF [3] and RecolorNeRF [2]. We include the complete metrics for all scenes of all three datasets in each table, one table for PSNR, one for SSIM and one for LPIPS. RecolorNeRF [2] did not converge to a solution in the Mip NeRF 360 dataset [1] so no numbers can be provided. Alongside the two baselines and our approach we include the Instant NGP [6] that was used as the initial pre-trained NeRF model. The performance of the base NeRF model conditions the performance of the recoloring methods and for such a reason we show its performance.

The results are grouped by metric in three different tables. In Table 1 the PSNR results for all scenes are showed, first the Synthetic NeRF [5] scene, second the LLFF [4] scenes and lastly the Mip NeRF 360 dataset [1] scenes. The scenes are named to make it clear to which dataset they belong and separated by a horizontal line. In Table 2 the results for the SSIM metric can be found and in Table 3 the results for the LPIPS metric are shown.

2. Additional Algorithm insights

Robust segmentation from a single view. A crucial aspect of our approach is its independence from additional multi-view segmentation inputs; instead, it leverages representations within NeRF. As shown in Fig. 1, the hashgrid features intrinsically learn semantic descriptions across multiple resolution levels. Features representing similar parts share similar values, maintaining consistency across various views. Notably, low resolution features remain similar within bigger elements (whole bonsai), while high resolution features capture smaller details (bonsai leaves). This combination of local and global effects is key to our method's ability to derive multiview segmentation masks from a single image.

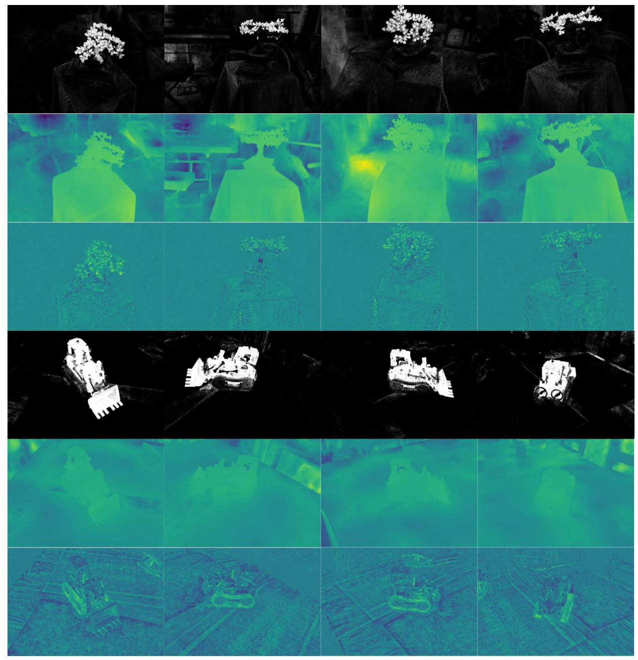

Figure 1. Row 1: Segmentation mask learned by IReNe, Row 2: Feature value at a low-res hashgrid level. Row 3: Feature value at a high-res hashgrid level. Top: bonsai scene; Bottom: lego scene.

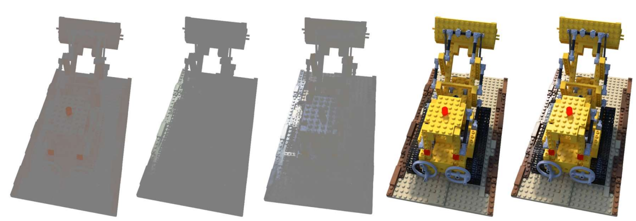

Figure 2. From left to right. Neuron render of: 1 diffuse; 1 view dependent (vd); all (47) vd; all (17) d; all neurons.

Capabilities and limitations of the method. Fine-tuning the final layer facilitates the application of global color adjustments. This is because the neurons in this layer essentially serve as color bases, with edits propagating from a single view. Fig. 2 shows a neuron functioning as a red color base and a different neuron preserving view-dependent effects. By doing neuron selection we preserve view-dependent effects, but it does not allow to modify them. However, both segmentation and neuron selection could potentially be adapted for texture editing purposes. To achieve this, modifications in the fine-tuning process of the last layer would be necessary.

Number of neurons selected in neuron selection. On average we select 16 neurons out of the possible 64. Out of the 6272 Color MLPs weights, we typically train only 48.

Dataset Scene	Baseline PSNR ↑	PaletteNeRF [3] PSNR ↑	RecolorNeRF [2] PSNR ↑	IReNe PSNR ↑
blender_drums	25.18	22.99	24.82	24.36
blender_ficus	32.37	30.36	34.54	31.69
blender_hotdog	36.31	24.23	22.70	27.70
blender_lego	34.32	33.91	31.92	34.39
blender_mic	36.24	31.61	31.06	33.52
llff_fern	25.07	21.33	26.43	25.34
llff_flower	29.12	27.30	27.79	28.58
llff_fortress	30.87	27.96	24.89	29.85
llff_horns	27.97	23.26	20.91	27.78
llff_leaves	17.44	17.08	17.62	17.38
llff_orchids	22.42	21.73	19.88	22.37
llff_room	26.85	23.52	28.89	25.94
llff_trex	27.89	27.32	26.14	27.80
mip360_bicycle	22.84	17.63	–	21.95
mip360_bonsai	31.19	27.33	–	30.31
mip360_counter	27.83	20.24	–	27.66
mip360_garden	25.41	23.91	–	25.24
mip360_kitchen	29.00	21.47	–	28.32
mip360_room	32.41	29.38	–	31.12
mip360_stump	22.59	18.24	–	22.99

Table 1. PSNR results of all baselines on all three datasets: Blender synthetic [5], LLFF [4], and Mip NeRF 360 [1]. We also provide the baseline metric value of the Instant NGP [6] that was used as the pre-trained NeRF and that gives a sense of the maximum performance that can be achieved. In the PSNR metric a higher value involves a better result.

Dataset Scene	Baseline SSIM ↑	PaletteNeRF [3] SSIM ↑	RecolorNeRF [2] SSIM ↑	IReNe SSIM ↑
blender_drums	0.93	0.91	0.93	0.92
blender_ficus	0.98	0.98	0.99	0.98
blender_hotdog	0.97	0.92	0.86	0.94
blender_lego	0.97	0.97	0.96	0.97
blender_mic	0.99	0.98	0.98	0.98
llff_fern	0.80	0.76	0.84	0.79
llff_flower	0.86	0.84	0.84	0.86
llff_fortress	0.84	0.82	0.61	0.84
llff_horns	0.85	0.81	0.61	0.84
llff_leaves	0.46	0.45	0.44	0.46
llff_orchids	0.73	0.73	0.58	0.73
llff_room	0.85	0.81	0.91	0.85
llff_trex	0.89	0.87	0.87	0.89
mip360_bicycle	0.51	0.46	–	0.49
mip360_bonsai	0.90	0.85	–	0.89
mip360_counter	0.80	0.72	–	0.80
mip360_garden	0.65	0.62	–	0.65
mip360_kitchen	0.79	0.72	–	0.79
mip360_room	0.90	0.88	–	0.90
mip360_stump	0.48	0.42	–	0.48

Table 2. SSIM results of all baselines on all three datasets: Blender synthetic [5], LLFF [4], and Mip NeRF 360 [1]. We also provide the baseline metric value of the Instant NGP [6] that was used as the pre-trained NeRF and that gives a sense of the maximum performance that can be achieved. In the SSIM metric a higher value involves a better result.

Dataset Scene	baseline LPIPS ↓	PaletteNeRF [3] LPIPS ↓	RecolorNeRF [2] LPIPS ↓	IReNe LPIPS ↓
blender_drums	0.06	0.09	0.06	0.08
blender_ficus	0.02	0.03	0.01	0.02
blender_hotdog	0.02	0.08	0.12	0.05
blender_lego	0.01	0.02	0.03	0.01
blender_mic	0.01	0.02	0.02	0.01
llff_fern	0.16	0.25	0.15	0.16
llff_flower	0.09	0.12	0.14	0.09
llff_fortress	0.10	0.14	0.57	0.11
llff_horns	0.12	0.20	0.59	0.13
llff_leaves	0.41	0.41	0.57	0.41
llff_orchids	0.15	0.16	0.35	0.15
llff_room	0.19	0.27	0.19	0.20
llff_trex	0.07	0.10	0.15	0.07
mip360_bicycle	0.48	0.62	–	0.54
mip360_bonsai	0.09	0.20	–	0.10
mip360_counter	0.18	0.22	–	0.18
mip360_garden	0.31	0.37	–	0.33
mip360_kitchen	0.16	0.29	–	0.17
mip360_room	0.12	0.15	–	0.13
mip360_stump	0.60	0.67	–	0.60

Table 3. LPIPS results of all baselines on all three datasets: Blender synthetic [5], LLFF [4], and Mip NeRF 360 [1]. We also provide the baseline metric value of the Instant NGP [6] that was used as the pre-trained NeRF and that gives a sense of the maximum performance that can be achieved. In the LPIPS metric a lower value involves a better result.

References

[1] Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022. 1, 3, 4, 5

[2] Bingchen Gong, Yuehao Wang, Xiaoguang Han, and Qi Dou. Recolornerf: Layer decomposed radiance fields for efficient color editing of 3d scenes. *arXiv preprint arXiv:2301.07958*, 2023. 1, 3, 4, 5

[3] Zhengfei Kuang, Fujun Luan, Sai Bi, Zhixin Shu, Gordon Wetzstein, and Kalyan Sunkavalli. Palettenerf: Palette-based appearance editing of neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023. 1, 3, 4, 5

[4] Ben Mildenhall, Pratul P. Srinivasan, Rodrigo Ortiz-Cayon, Nima Khademi Kalantari, Ravi Ramamoorthi, Ren Ng, and Abhishek Kar. Local light field fusion: Practical view synthesis with prescriptive sampling guidelines. *ACM Transactions on Graphics (TOG)*, 2019. 1, 3, 4, 5

[5] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2020. 1, 3, 4, 5

[6] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.*, July 2022. 1, 3, 4, 5